into the process as appropriate. **RESULTS:** Two main linguistic and conceptual issues emerged during the translation process. Firstly, when an original item used more than one adjective to cover a single concept, some languages only had one term to express this. Secondly, there was the challenge of using culturally appropriate expressions for taboo concepts such as suicide and self-harm, to ensure homogenous response across all languages. **CONCLUSIONS:** The language versions of the GAD-7 and PHQ-9 were established according to a rigorous standardized methodology to facilitate international comparison and pooling of data. The linguistic validation process aims to ensure conceptual equivalence across different language versions on the basis of a pre-defined concept list explaining what the original instrument should measure. The process as a whole supports the advantage of integrating international feedback on concepts and wording during the development of questionnaires.

**PMC73**

## EVALUATION OF GLP-1 PRODUCT ATTRIBUTES IN TREATING PEOPLE WITH TYPE-2 DIABETES IN US: COMPARING TIME-TRADEOFF AND WILLINGNESS-TO-PAY METHODOLOGIES

Zanutto E[1], Conner C[2], Polster M[3], McDonald SS[1], Hammer M[4]
[1]National Analysts Worldwide, Philadelphia, PA, USA, [2]Novo Nordisk, Inc., Princeton, NJ, USA, [3]National Analysts Worldwide, philadelphia, PA, USA, [4]Novo Nordisk A/S, Bagsværd, Denmark

**OBJECTIVES:** To assess patient utilities for cost utility analyses in health economics two different ways of revealing preferences have been compared. Time trade-off (TTO) and willingness-to-pay (WTP) surveys have been completed comparing patients' preference in product propositions. TTO yields estimates of the amount of time that patients would be willing forego to achieve preferred products attributes, and WTP yields estimates of the amount of money that patients are willing to pay for products attributes. Despite their widespread use, results of the two methodologies have not, to our knowledge, been compared directly in diabetes. **METHODS:** The two methods were used to evaluate the reactions of people with type-2 diabetes to two GLP-1 injectable diabetes medications that varied on four attributes: efficacy in controlling blood glucose (measured by HbA1c), dosing frequency (twice-daily vs. once-daily), incidence of hypoglycemia, and incidence of nausea. To maximize statistical power and allow comparisons across patient groups, a large internet-based survey (more than 500 respondents) was conducted in U.S. with four categories of self-identified patients who were sampled based on their medication history, and randomly assigned to either the TTO or WTP. **RESULTS:** Results suggest that the WTP methodology yields greater face validity and sensitivity than TTO (100% of respondents prefer the superior profile in WTP vs. 96% in TTO). Data from conjoint analysis designed to establish the importance of each of the four product attributes in the decision-making process used by patients reveal similar patterns of results for the two methodologies. **CONCLUSIONS:** Regardless of whether patients were in the group assigned to TTO or WTP, patients perceive efficacy (HbA1c control) to be the most important product attribute, followed by incidence of nausea. Patients evaluated the incidence of hypoglycemia and dosing schedule as less important relative to the HbA1c control and nausea in the decision-making process.

**PMC74**

## QUALITY-ADJUSTED LIFE YEARS SAVED BY PREVENTION OF HEAD INJURY THROUGH ENFORCEMENT OF HELMET LAW

Lee HY[1], Chen YH[2], Wang JD[3]
[1]National Taiwan University Hospital, Taipei, Taiwan, [2]University of Michigan, Ann Arbor, MI, USA, [3]National Taiwan University, College of Public Health, Taipei, Taiwan

**OBJECTIVES:** To evaluate the potential long-term health impact of helmet law, we calculated the loss of quality-adjusted life year (QALY) under different proportions of helmet wearing among the motorcyclists. **METHODS:** The quality-adjusted life expectancies for helmeted and non-helmeted motorcyclists were estimated by adjusting the survival function based upon the Head Injury Registry with quality of life measures assessed under the EQ-5D questionnaire. We took Hualien County, where a lower rate of helmet wearing (77%) was reported, as an example to calculate the expected numbers of prevented head-injured cases by multiplying the population at risk with the incidence rate of head injury for helmeted and non-helmeted motorcyclists. As different proportions of helmet-wearing and different proportions of full face helmet in motorcyclists were assumed, the expected numbers of prevented cases were calculated, which were multiplied with the loss of QALE of an average case to predict the potential benefit of helmet use. **RESULTS:** Under the current proportion of helmet wearing, the annual loss associated with head injured was 4346.9 QALYs in Hualien County. If the proportion of helmet wearing could be increased to 100%, the health benefit saved was estimated 1434.3 QALYs. If 80% of them used full face helmet, the total gain was increased to 2500 QALYs. **CONCLUSIONS:** The health benefit of helmet protection for head injury can be determined under units of quality-adjusted life year (QALY) and directly applied in future cost-effective analyses for public health policy.

**PMC75**

## MEANINGFUL VERSUS USEABLE RESPONSES TO PREFERENCE SURVEYS: INSIGHTS INTO IMPROVING THE VALIDITY OF HEALTH UTILITY SCORES

Wittenberg E[1], Prosser LA[2]
[1]Brandeis University, Waltham, MA, USA, [2]University of Michigan, Ann Arbor, MI, USA

**OBJECTIVES:** To identify the types of and reasons for unusable preference survey responses with the goal of informing improvements in health utility assessment methodology. **METHODS:** We define typologies that represent unusable responses from health utility surveys; present evidence from the literature on the frequency of such responses; present empiric data on the rationale for such types of responses; and discuss methods for handing data that contain such responses, and implications for interpreting analytic results based on health utility data. **RESULTS:** Potentially unusable health utility survey responses include (1) illogical, (2) inconsistent, (3) invariant, and (4) "protest" responses, plus (5) refusals. These responses may represent simple mistakes or misunderstandings of the survey task, which introduce noise into results, or they may be intentional responses to the parameters of the survey task that may confound respondents' other values with the value of the health state being assessed. Unusable responses can be avoided through anticipation and careful design of survey instruments, particularly for specific populations and health states, including cognitive testing prior to fielding. Unusable responses can also be omitted from analyses or analyzed separately. **CONCLUSIONS:** Unusable health utility survey responses challenge the validity of utility estimates and all analyses that incorporate these values, so it is critical to minimize these responses. Mechanisms to correct errors are useful, but may not address true preferences that are in response to elements of the measurement task and hence will not be corrected. Correction mechanisms may include tailoring the task to particular situations when bias is anticipated, such as parent valuations of children's states or individuals who express religious beliefs. Recognition of the prevalence of unusable data in health utility data sets and identifying methods for handling these errors is essential to understand the accuracy and precision of results and analyses that depend thereon.

**PMC76**

## AN ASSESSMENT OF THE EVIDENCE TO SUPPORT THE USE OF LINGUISTIC VALIDATION INTERVIEWS AS AN INTEGRAL PART OF THE TRANSLATION OF PROS FOR USE IN MULTI-NATIONAL TRIALS

Gordon-Stables R, Houchin C, Wild D
Oxford Outcomes Ltd, Oxford, UK

Linguistic validation interviews are considered a vital step in the generally accepted methodology for the translation of PRO measures. The translated PRO measure is administered to 5 or more native-speakers of the target language who also have a diagnosis of the condition of interest. **OBJECTIVES:** The objective of this study was to determine whether there is any evidence to suggest that the changes made are significant enough to make the step worthwhile. **METHODS:** Examples of linguistic validation changes were gathered from completed PRO translation projects. For one scale for which 26 translations were developed, changes made at the linguistic validation stage were examined quantitatively and qualitatively in order to determine the number of changes made and the nature of these changes. The number of changes made per language as a result of the linguistic validation step ranged from 0 to 11, with a mean of 1.4 changes. The changes made can be categorised as follows: Items where participants reported they did not understand. Items where participants reported that they understood, but probing made it clear that confusion had arisen. Items where participants picked up on vocabulary that was not what they would usually use, and items where participants spotted spelling and grammatical errors. **RESULTS:** Whilst in purely numerical terms, the contribution of linguistic validation to the translation process may seem small; in terms of the importance of those changes and in the confirmation of accuracy of the translated concept, its contribution is great. Whilst the translation may appear to be "correct" in other stages, only linguistic validation interviews show if the target population understands the translation as intended. **CONCLUSIONS:** The goal of translation of a PRO instrument is to achieve conceptual equivalence, and linguistic validation is arguably the most important step in assuring this.

**PMC77**

## DEVELOPMENT AND VALIDATION OF AN E-PRO TEMPLATE ADMINISTERED VIA A WEB-BASED INTERFACE

Eremenco S[1], Coyne KS[1], Duran B[2], Leidy NK[1]
[1]United BioSource Corporation, Bethesda, MD, USA, [2]United BioSource Corporation, Lexington, MA, USA

BACKGROUND: As electronic patient-reported outcome (ePRO) use expands, questions have arisen regarding the effect of presentation form and style on patient interpretation and whether there is an optimal format to ensure clarity, ease of use, and data quality. **OBJECTIVES:** To develop and validate a standardized ePRO interface of common PRO question formats for use on desktop and tablet PCs to streamline ePRO study implementation. **METHODS:** A web-based application was developed for desktop and tablet platforms with 52 demographic and clinical questions representing Likert-type, dichotomous, VAS and numeric rating scales. Adults from the general population were recruited through newspaper advertisements to participate in questionnaire completion and a cognitive debriefing interview. Participants answered questions on either desktop or tablet PC and then reviewed the alternate format during the interview. Alternative question formats were reviewed to obtain further respondent insight and recommendations. **RESULTS:** Participants: N = 47: mean age 54 years (23 to 79), 51% female; 49% reported a chronic condition. 13%<= igh school education. Average completion time: desktop (n = 22) = 11 minutes; tablet PC (n = 25) = 13 minutes, with all participants finding their completion time acceptable. A total of 53% preferred desktop to tablet. Respondents found question formats easy to answer, except items requiring date and time entry (resolved through minor instruction revisions) and the VAS. The VAS was presented three ways: anchors at 0 and 100, hashmarks every ten digits, and hashmarks only with a display of the number selected. 77% preferred hashmarks; 83% liked the number displayed. No participants changed their answer after the number appeared. **CONCLUSIONS:** Respondents provided useful information on their perception and preferences for question format presented

on desktop and tablet PC platforms. Results of the interviews informed the development and final testing of a standardized interface that can be used across studies to facilitate rapid deployment of ePRO studies.

**PMC79**

### TRANSLATION OF THE PATIENT-REPORTED OUTCOMES MEASUREMENT INFORMATION SYSTEM INTO SPANISH

Arnold BJ, Correia H, Cella D
NorthShore University HealthSystem, Evanston, IL, USA

The Patient-Reported Outcomes Measurement Information System (PROMIS) provides accurate and efficient measurement of patient-reported outcomes. Developed in English using qualitative methods, PROMIS seeks to measure symptoms, such as pain and fatigue, and aspects of health-related quality of life across a wide variety of chronic diseases and conditions. **OBJECTIVES:** In order to enable participation of the rapidly growing Spanish-speaking population of the USA it was necessary to translate PROMIS banks from English into Spanish using methods that would ensure linguistic equivalence and cultural appropriateness. **METHODS:** Five hundred and twenty-two items were translated into Spanish using the FACIT Multilingual Translation Methodology which consists of the following twelve steps: (1) translatability review of existing English items, (2) creation of item definitions, (3) two simultaneous forward translations,(4) reconciliation of forward translations, (5) back translation of reconciliation, (6) expert review of back translation and previous steps, (7) preliminary finalization for pilot-testing, (8) harmonization, (9) quality assurance, (10) formatting, (11) cognitive testing with native speakers of Spanish and (12) analysis and finalization of translations based on qualitative data collected during pilot-testing. Recognizing the need to address diversity within the Spanish-speaking population of the USA, linguists from various Spanish-speaking regions across the globe were recruited to achieve a universal Spanish translation. **RESULTS:** Some of the linguistic challenges encountered during the translation process as well as the language solutions for resolving difficulties associated with the cultural and linguistic heterogeneity of the Spanish-speaking population residing in the USA will be highlighted in this presentation. **CONCLUSIONS:** Future research includes further validation of the Spanish translations using psychometric testing of the equivalence of banks in English and Spanish, including assessment of differential item functioning across different cultural groups. The translation of additional items into Spanish and into other languages is also explored in this presentation.

**PMC81**

### ELECTRONIC PATIENT REPORTED OUTCOMES: FOLLOWING FDA GUIDANCE FROM A VENDOR PERSPECTIVE

Ross J, Shea E
Almac Clinical Technologies, Yardley, PA, USA

**OBJECTIVES:** To provide an overview of considerations sponsors/vendors need to address in order to meet FDA expectations during the planning and implementation phases when using Electronic Patient-Reported Outcomes (ePROs) in clinical trials, to understand important considerations required in ePRO-use when planning and implementing clinical trials, identify essential considerations from a vendor's perspective, and following FDA guidance as a vendor or when using a vendor. **METHODS:** Sponsors often utilize vendors for the planning and implementation of their trials. When vendors are involved in these phases with ePRO, it is not only the sponsor's responsibility, but also the vendor's to address considerations to ensure FDA expectations are met. However, meeting these expectations can be challenging across the pharmaceutical industry. To assist industry, the FDA released a Draft Guidance to communicate their perspective on how they evaluate ePRO-use for efficacy endpoints in clinical trials and for support claims made in approved product labeling. **RESULTS:** Often, sponsors/vendors are unsure how to follow the guidance and are unaware of important considerations necessary when incorporating them in clinical trials. These considerations can include: standardizing data collection with electronic tools, handling missing values, validation, reliability, and responsiveness to clinically significant differences. EPRO-use in clinical trials requires careful planning and execution. When these considerations are not satisfied, the trials can face serious consequences by the FDA throughout the product development and approval processes. **CONCLUSIONS:** This session intends to provide an overview of how FDA ePRO-use expectations should be met from the perspective of a vendor during the planning and implementation phases of clinical trials. Key considerations will be discussed. Fictitious examples of what could go wrong will be presented. A summary of recommendations will be provided on how to follow the FDA guidance and avoid making critical errors when employing ePROs in clinical trials.

**PMC82**

### PSYCHOMETRIC PROPERTIES OF FOUR KISWAHILI TRANSLATED DISEASE SPECIFIC PATIENT REPORTED OUTCOME MEASURES

Kangethe AW, Franic DM
University of Georgia, Athens, GA, USA

**OBJECTIVES:** Kiswahili (Swahili) is the operative language of 100 million East Africans. Most patient reported outcome (PRO) measures are developed in the Anglo-American literature; however, no disease specific studies have been published on their translation into Kiswahili. This project assessed the 1) psychometric properties, and 2) cross cultural adaptation of disease specific (PRO) measures translated from English .to Kiswahili. **METHODS:** A comprehensive literature and web search identified four Kiswahili translated PRO measures for study inclusion. PRO measures were included for evaluation: if there was at least one peer reviewed English publication providing psychometric data, and instrument availability. Psychometric instrument evaluation criteria were based on: conceptual model, practicality (5–15 minutes), depth, reliability (internal consistency, test-retest: greater than or equal to .7), validity (convergent and divergent), and responsiveness (Lohr, 2002). Assessment of the procedures used to evaluate conceptual and linguistic equivalence was based on recommendations by ISPOR Taskforce for Translation and Cultural Adaptation (Wild et al, 2005). **RESULTS:** Of the four instruments evaluated: Oral Impacts on Daily Performance (OIDP), Patient Health Questionnaire – 9 (PHQ-9), World Health Organization Prevention of Blindness and Deafness Visual Functioning-20 (WHO/PBD VF20) and Hopkins Symptom Check List-25 (HSCL-25), none met all study criteria for psychometric properties or linguistic adaptation procedures. Conceptual model, practicality and internal consistency were met for all scales. In cases where criteria were not met it was predominantly due to missing data, e.g, divergent validity was unavailable for the PROs assessed. **CONCLUSIONS:** Although attempts to provide good practice guidelines on translation of PRO instruments have been made, there is need to reach a consensus on these guidelines. Researchers involved in the cross cultural adaptation of PRO instruments are encouraged to be transparent to enable the complete evaluation and subsequent adoption of the translated scale into clinical practice or clinical trials.

**PMC83**

### USE OF CONJOINT ANALYSIS IN HEALTH OUTCOMES RESEARCH: AN EXAMINATION OF THE LITERATURE

Szeinbach SL, Beyer AP, Qureshi ZP, Uhas AA, Visaria J, Seoane-Vazquez E
Ohio State University, Columbus, OH, USA

Patient-reported outcomes (PRO) are key endpoint measures to examine and assess preferences for health improvements. Conjoint analysis (CA), traditionally used in marketing studies to examine tradeoffs among attributes, has gained popularity in health care and PRO studies. However, the use of CA and the type of CA studies performed in health care has not been fully characterized. In this study, we characterized the current trends for CA, classified CA studies by area of focus, examined study complexity, and provided recommendations for future research. Literature reviews were conducted using a multi-database search from 2000 to 2008. Review articles, methodological, and non-health related CA studies were excluded. Preference studies were coupled with key words such as: conjoint, health, disease, evaluation, discrete choice and outcome(s). Five years were selected for detailed content analysis and rechecked for accuracy. For the five years examined, there was an upward trend for the number of health-related CA studies (2000 = 22; 2002 = 39; 2004 = 38; 2006 = 56; 2008 = 70). However, results from Chi-square analysis revealed no significant differences among years for the general area of focus for CA studies performed. The largest proportion of CA studies consisted of attribute importance, risk, and pharmacoeconomics studies where utility estimates were used to assess willingness-to-pay, quality of life, or quality-adjusted life years. Other applications included disease screening, value of services, satisfaction, treatment evaluation and service delivery. The use of CA to assess PROs in health research has expanded. Recent and more innovative applications have extended to adherence, disease screening, technology, and value. More research is needed to evaluate the usefulness of CA for large database studies and for economic analyses in health technology assessment.

**PMC84**

### THE POWER OF ASSUMPTIONS

van Hout BA
Pharmerit Ltd, York, North Yorkshire, UK

**OBJECTIVES:** Studies powered using a dichotomous endpoint, are often too small to find significant differences in quality of life (QoL) or costs. Including the likelihood that events in both arms are similar, using either frequentist assumptions or Bayesian priors, may increase the power. **METHODS:** A study of patients with late pain is imagined. With therapy, 92% is expected to be pain-free, without 85%. With $2 \times 250$ patients, the dichotomous power (pain: yes/no) is 80%. Using the EQ-5D utility score it is 40% (with the EQ-5D-pain-dimension at 3 with pain and at 0 without pain and the other EQ-5D-dimensions at random population levels). Trials are simulated and T-tests are calculated based on: 1. QoL of patients with pain is identical in both arms; 2. QoL of patients without pain is identical in both arms; 3. 1 + 2. 95% credibility intervals are calculated using normal priors concerning the difference in QOL (per arm) with and without pain. Expectations and precisions are varied as well as base line probabilities. **RESULTS:** Making both assumptions, using T-tests, increases the power from 40% to 80%. Assumption 1 does so by 2%, assumption 2 by 79%. Both assumptions contribute equally when the expected pain-free levels are approximately 55% versus 44% instead of 95% versus 82%. The Bayesian model coincides with the frequentist approach when the precision in the priors concerning the differences in QoL are set to the extremes (zero or infinity). Between the extremes the Bayesian approach offers the flexibility to compromise. The power increase between the extremes can be characterized by a cumulative normal distributions on the log of the squared root of the precisions. **CONCLUSIONS:** Predefining logical assumptions in a QoL analysis may increase the power of a study. The larger the group of patients the assumption is applied to, the bigger the power increase.